

Discovering biomarkers associated and predicting cardiovascular disease with high accuracy using a novel nexus of machine learning techniques for precision medicine

Year: 2023 | Citations: 514 | Authors: William DeGroat, Habiba Abdelhalim, K. Patel, Dinesh Mendhe, Saman Zeeshan

Abstract

Personalized interventions are deemed vital given the intricate characteristics, advancement, inherent genetic composition, and diversity of cardiovascular diseases (CVDs). The appropriate utilization of artificial intelligence (AI) and machine learning (ML) methodologies can yield novel understandings of CVDs, enabling improved personalized treatments through predictive analysis and deep phenotyping. In this study, we proposed and employed a novel approach combining traditional statistics and a nexus of cutting-edge AI/ML techniques to identify significant biomarkers for our predictive engine by analyzing the complete transcriptome of CVD patients. After robust gene expression data pre-processing, we utilized three statistical tests (Pearson correlation, Chi-square test, and ANOVA) to assess the differences in transcriptomic expression and clinical characteristics between healthy individuals and CVD patients. Next, the recursive feature elimination classifier assigned rankings to transcriptomic features based on their relation to the case–control variable. The top ten percent of commonly observed significant biomarkers were evaluated using four unique ML classifiers (Random Forest, Support Vector Machine, Xtreme Gradient Boosting Decision Trees, and k-Nearest Neighbors). After optimizing hyperparameters, the ensembled models, which were implemented using a soft voting classifier, accurately differentiated between patients and healthy individuals. We have uncovered 18 transcriptomic biomarkers that are highly significant in the CVD population that were used to predict disease with up to 96% accuracy. Additionally, we cross-validated our results with clinical records collected from patients in our cohort. The identified biomarkers served as potential indicators for early detection of CVDs. With its successful implementation, our newly developed predictive engine provides a valuable framework for identifying patients with CVDs based on their biomarker profiles.